The Role of the UMLS in Vocabulary Control

CENDI Conference

"Controlled Vocabulary and the Internet"

Stuart J. Nelson, MD

Observations

- Words are not enough
- Word based synonymy is not enough
- Single phrases are not enough
- Need “web-scale” synonymy

Synonymy that is “Web-Scale”

- Concepts (classes of terms) finely granular
- “Fully expressive” names
  - Acronyms
  - Gene Names
  - “special meanings”
- Scalable methodologies, large scale vocabularies

UMLS Purpose

- Make it easy for health professionals and researchers to retrieve and integrate relevant information from disparate automated sources, e.g.
  - computer-based patient records
  - factual databanks
  - bibliographic databases and full-text
  - expert systems
- Antedated and anticipated the Web

UMLS Focus
Conceptual Connections

- Build knowledge sources that can be used by intelligent programs to overcome:
  - disparities in language used by different users and in different information sources;
  - difficulties in identifying which of many information sources is relevant

UMLS Knowledge Sources

Multi-purpose tools or “intellectual middleware” for System Developers

- Metathesaurus
- SPECIALIST lexicon and lexical programs
- Semantic Network

UMLS Knowledge Sources Distribution

- Annual updates, 1990 - -
- Free under license agreement with NLM
  - Need separate license agreements with vocabulary producers for some uses of some vocabularies in the Metathesaurus
- Available to licensed users (~900) via Internet server and on CDs

1999 UMLS Metathesaurus

- 626,313 concepts (Oculus, eye =1)
- 1,134,413 “terms” (Eye, Eyes, eye = 1)
- 1,358,891 “strings”/concept names
  - (Eye, Eyes, eye = 3)
- ~50 source vocabularies

UMLS Metathesaurus *Finely Granular Concepts*

- Concepts, terms, and attributes from many controlled vocabularies
- New inter-source relationships, definitional information, use information
- Scope determined by combined scope of source vocabularies
- Strict definition of synonymy
- Semantic neighborhood

UMLS Source Vocabularies

- Widely varying purposes, structures, properties
  - Thesauri, e.g., MeSH
  - Statistical Classifications, e.g., ICD
  - Billing Codes, e.g., CPT
  - Clinical coding systems, e.g., SNOMED
  - Lists of controlled terms, e.g., COSTAR, HL7 value sets

Metathesaurus Construction

The Scalable Methodology

- Convert machine-readable vocabulary sources to UMLS “normal” form, making source semantics explicit
- Merge, using source semantics and lexical processing techniques
- Edit results, adding additional relationships and semantic information

Metathesaurus Characteristics (1)

- Concept organization
- Many sources in a common database format
- Representation of the meaning in each source vocabulary
- Explicit tagging of each source vocabulary's information

Current MeSH -- Organized by Preferred Term

D0015154

Esophageal Motility Disorders (MH)

- **Esophageal Dysmotility (EP - SYN)**
- **Nutcracker Esophagus (EP - NRW)**

UMLS Metathesaurus -- Organized by Concept

C0014858

Esophageal Motility Disorders (MeSH, Read)

Esophageal Dysmotility (MeSH, Read)

Oesphageal Dysmotility (Read)

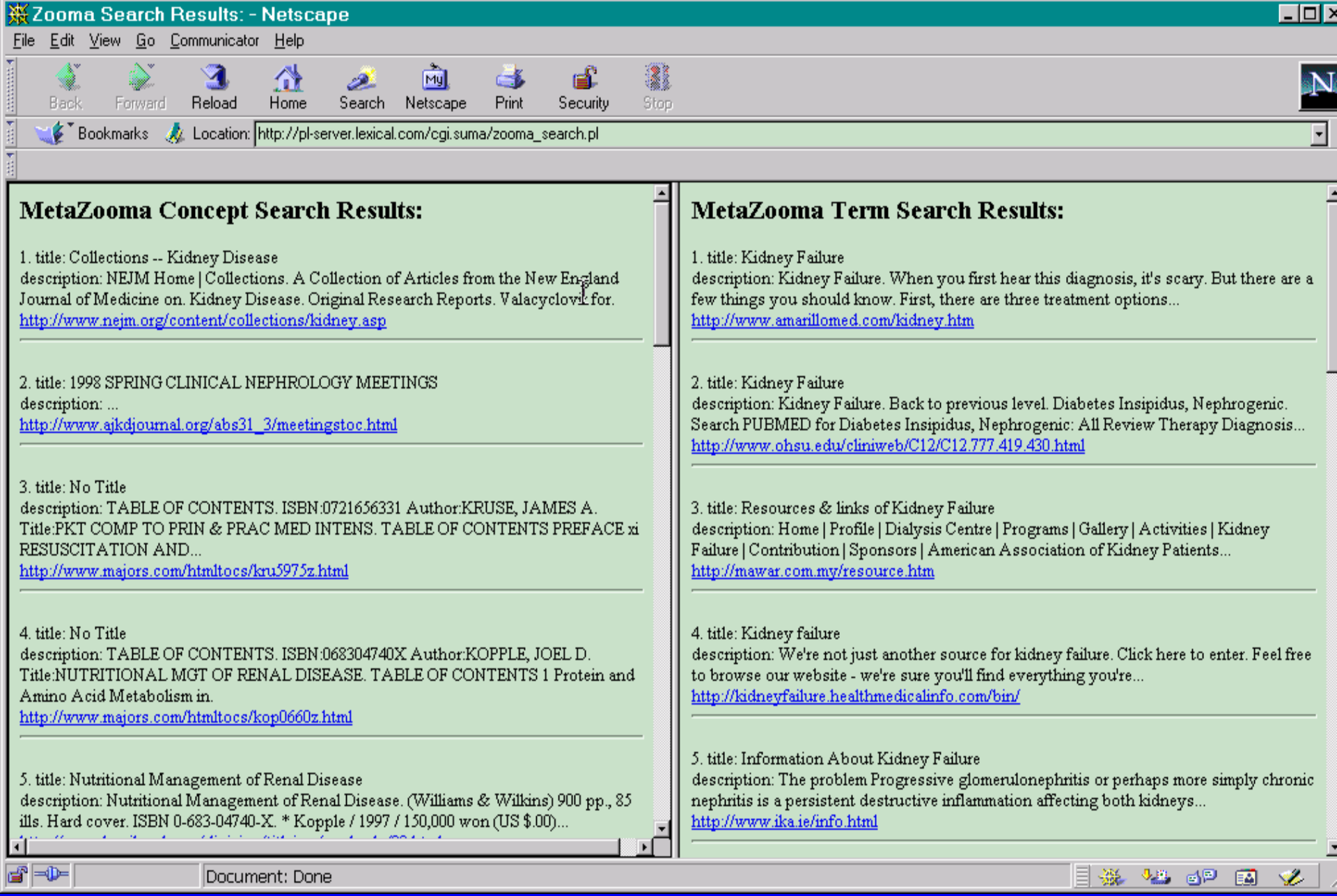
C0028705

Nutcracker Esophagus (MeSH, Read)

Symptomatic esophageal peristalsis (Read)

Websites Using UMLS

- Medical World Search -
- CliniWeb - OHSU
- MetaZoomA -Lexical/Apelon


Zooma Search Results: - Netscape
File Edit View Go Communicator Help
Back Forward Reload Home Search Netscape Print Security Stop
Bookmarks Location: http://pl-server.lexical.com/cgi.suma/zooma_search.pl
MetaZooma Concept Search Results:
1. title: Collections -- Kidney Disease
description: NEJM Home | Collections. A Collection of Articles from the New England Journal of Medicine on. Kidney Disease. Original Research Reports. Valacyclovir for.
http://www.nejm.org/content/collections/kidney.asp
2. title: 1998 SPRING CLINICAL NEPHROLOGY MEETINGS
description: ...
http://www.ajkdjournal.org/abs31_3/meetingstoc.html
3. title: No Title
description: TABLE OF CONTENTS. ISBN:0721656331 Author:KRUSE, JAMES A. Title:PKT COMP TO PRIN & PRAC MED INTENS. TABLE OF CONTENTS PREFACE xi RESUSCITATION AND...
http://www.majors.com/htmltocs/kru5975z.html
4. title: No Title
description: TABLE OF CONTENTS. ISBN:068304740X Author:KOPPLE, JOEL D. Title:NUTRITIONAL MGT OF RENAL DISEASE. TABLE OF CONTENTS 1 Protein and Amino Acid Metabolism in.
http://www.majors.com/htmltocs/kop0660z.html
5. title: Nutritional Management of Renal Disease
description: Nutritional Management of Renal Disease. (Williams & Wilkins) 900 pp., 85 ills. Hard cover. ISBN 0-683-04740-X. * Kopple / 1997 / 150,000 won (US $.00)...
MetaZooma Term Search Results:
1. title: Kidney Failure
description: Kidney Failure. When you first hear this diagnosis, it's scary. But there are a few things you should know. First, there are three treatment options...
http://www.amarillomed.com/kidney.htm
2. title: Kidney Failure
description: Kidney Failure. Back to previous level. Diabetes Insipidus, Nephrogenic. Search PUBMED for Diabetes Insipidus, Nephrogenic: All Review Therapy Diagnosis...
http://www.ohsu.edu/cliniweb/C12/C12.777.419.430.html
3. title: Resources & links of Kidney Failure
description: Home | Profile | Dialysis Centre | Programs | Gallery | Activities | Kidney Failure | Contribution | Sponsors | American Association of Kidney Patients...
http://mawar.com.my/resource.htm
4. title: Kidney failure
description: We're not just another source for kidney failure. Click here to enter. Feel free to browse our website - we're sure you'll find everything you're...
http://kidneyfailure.healthmedicalinfo.com/bin/
5. title: Information About Kidney Failure
description: The problem Progressive glomerulonephritis or perhaps more simply chronic nephritis is a persistent destructive inflammation affecting both kidneys...
http://www.ika.ie/info.html
Document: Done

Cautions in Searching

- Mention is not Aboutness, but . . .
- Aboutness is not Relevance
- Relevance is in the eye of the beholder

UMLS Summary

- Concept-based
- Extensive content in biomedicine
- Scalable methodology
- Supporting both retrieval whether indexed or searching full-text

Unified Medical Language System (UMLS) - Netscape
File Edit View Go Communicator Help
Bookmarks Location: http://www.nlm.nih.gov/research/umls/umlsmain.html
Instant Message Internet Lookup New&Cool

Unified Medical Language System (UMLS)

NLM's Unified Medical Language System (UMLS) project develops and distributes multi-purpose, electronic "Knowledge Sources" and associated lexical programs. System developers can use the UMLS products to enhance their applications -- in systems focused on patient data, digital libraries, Web and bibliographic retrieval, natural language processing, and decision support. Researchers will find the UMLS products useful in investigating knowledge representation and retrieval questions.

- UMLS Knowledge Source Server is available to those who have signed the UMLS license agreement.
- License Agreement for use of the UMLS Knowledge Sources includes a list of vocabularies in the UMLS Metathesaurus. The UMLS products are available free of charge to U.S. and international users. Use of the UMLS Metathesaurus may require additional agreements (which may involve fees) with producers of the individual vocabularies it contains.
- Obtaining Access to UMLS Resources
- UMLS Applications
- Unified Medical Language System Fact Sheet
  - UMLS Metathesaurus Fact Sheet
  - SPECIALIST Lexicon Fact Sheet
  - UMLS Semantic Network Fact Sheet
- UMLS Documentation
  contains complete description of the Knowledge Sources and their distribution formats.
- Comprehensive Bibliography 1986-96
  For more recent articles search Unified Medical Language System in MEDLINE.

Send questions, comments about the UMLS project to: custserv@nlm.nih.gov or call 1-888-FINDNLM.

U.S. National Library of Medicine (NLM)

Document: Done